Multi-Mean Reverting Processes : Analytical and Statistical Approaches

Benoît Nieto (CMAP, Polytechnique)

under the supervision of :

Christophette Blanchet-Scalliet (ECL, ICJ) and Diana Dorobantu (ISFA)

Introduction

General model

The process $(X_t)_{t\geq 0}$ is solution to SDE:

$$X_t = X_0 + \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dB_s, \quad t \geq 0, \tag{1}$$

where $B = (B_t)_{t\geq 0}$ is a Brownian motion and the functions μ and σ possibly exhibit discontinuities.

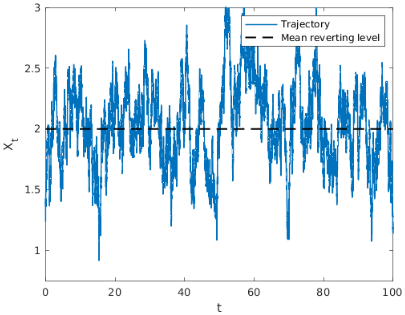
A threshold process is a stochastic process where the dynamics change when the process crosses specific threshold levels.

Example: $\sigma(x) = \mathbb{1}_{x<0} + 2\mathbb{1}_{x\geq 0}$.

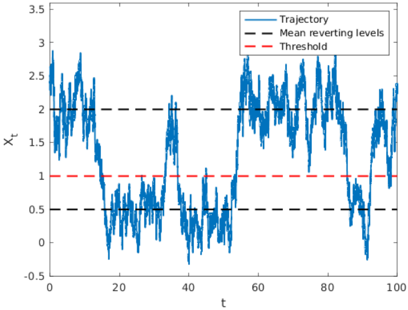
A multi mean-reverting process is a stochastic process where trajectories tend to revert towards multiple long-term average values, depending on the state or regime of the process.

Example: $\mu(x) = 1 - x$, $\sigma(x) = 1$, one mean reversion level equal to 1.

Illustration



→ One mean reversion level



→ One threshold and two mean reversion levels

Example : Ornstein-Uhlenbeck (OU) process, Cox-Ingeroll Ross (CIR) process ...

Example : Threshold OU process, Threshold CIR process, Threshold CKLS process ...

Thesis Outline :

1. **Part 1** :

   - A pseudo-likelihood estimator of the Ornstein-Uhlenbeck parameters from suprema observations.

     Blanchet-Scalliet C., Dorobantu D., Nieto B. (2024). A pseudo-likelihood estimator of the Ornstein-Uhlenbeck parameters from suprema observations. *Statistical Inference for Stochastic Processes.* Volume 27, pages 407–425.

   - Some properties for μ-zeros of Parabolic Cylinder functions.

     Blanchet-Scalliet C., Dorobantu D., Nieto B. (2023). Some properties for ν-zeros of Parabolic Cylinder functions. *Le Matematiche.* Volume LXXVIII, pages 277-287.

2. **Part 2** :

   - The killed Threshold Ornstein-Uhlenbeck process.

   - Parameters estimation of a Threshold CKLS process from continuous and discrete observations.

     Mazzonetto S., Nieto B. (2024). Parameters estimation of a Threshold CKLS process from continuous and discrete observations. *Submitted.*

Summary

1. A pseudo-likelihood estimator of the Ornstein-Uhlenbeck parameters from suprema observations

2. Parameters estimation of a Threshold CKLS process from continuous and discrete observations

1. A pseudo-likelihood estimator of the Ornstein-Uhlenbeck parameters from suprema observations

Introduction

The Ornstein-Uhlenbeck (OU) process is a stochastic process that exhibits mean-reverting behaviour.

Example:

- In Finance, the OU is used in the Vasicek model to describe the evolution of interest rates.
  - Estimation using trajectory observation the process.

Franco J.C.G. (2003). Maximum likelihood estimation of mean reverting processes *Real Options Practice*

- The OU process can model for the spontaneous activity of a neuron.
  - Estimation using first hitting times observations.

Mullowney P. and Iyengar S. (2008). Parameter estimation for a leaky integrate-and- fire neuronal model from ISI data *Journal of Computational Neuroscience*

Temperature Dynamic

- Temperature dynamic modelled by a mean-reverting process such as an Ornstein-Uhlenbeck (OU) process.

Alaton P., Djehiche B. and Stillberger D. (2002). On modelling and pricing weather derivatives *Energy & Power Risk Management*

Dischel B. (1998). At last : A model for weather risk *Applied Mathematical Finance*

- Purpose : estimate the parameters of this process thanks to the daily suprema temperatures.

- Why ? forecasting and assessing risk measures such as the probability of heat wave in summer (to be over a threshold during a certain period).

Content

Context

Estimate the drift and volatility parameters by using **suprema observations** of a stationary OU process.

[Blanchet-Scalliet & al, 2018]

Method: Least square method based on quantiles.

$\rightarrow$ Theoritically: no statistical properties on the estimator.

$\rightarrow$ Numerically: the method is computationally expensive.

Our goal

Method: Pseudo-likelihood.

$\rightarrow$ Theoritically: study of the asymptotic behavior of the estimator.

$\rightarrow$ Numerically: develop a method with low numerical cost and better accuracy.

Model

Ornstein-Uhlenbeck (OU) Process

The process $(X_t)_{t\geq 0}$ is the solution of the following SDE :

$$dX_t = (a - bX_t)dt + \sqrt{\beta} dB_t, \quad X_0 \underset{law}{\sim} \mathcal{N}\left(\frac{a}{b}, \frac{\beta}{2b}\right), \tag{2}$$

with $b, \beta \in \mathbb{R}_+^*$, $a \in \mathbb{R}$ and $(B_t)_{t\geq 0}$ a standard Brownian motion, X_0 independent of $(B_t)_{t\geq 0}$.

We denote $\theta = (a, b, \beta) \in \Theta$, with Θ a compact subset of $\mathbb{R} \times \mathbb{R}_+^\star \times \mathbb{R}_+^\star$.

Property

- The process $(X_t)_{t\geq 0}$ is a Markov process.
- As b is strictly positive, $(X_t)_{t\geq 0}$ is ergodic.
- As $X_0 \sim \mathcal{N}\left(\frac{a}{b}, \frac{\beta}{2b}\right)$, $(X_t)_{t\geq 0}$ is a stationary process.

Let us note $(S^{i,0})_{i\geq 1}$, the following sequence:

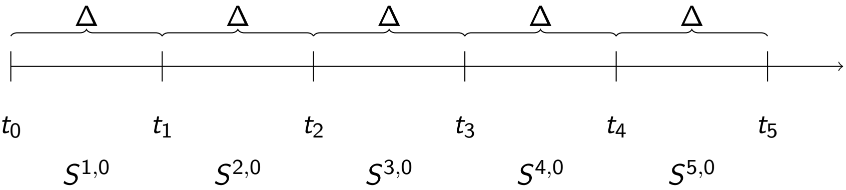
$$S^{i,0} = \sup_{s\in[t_{i-1},t_i]} X_s,$$

with $(t_i)_{i\geq 0}$ a time sequence such that $t_0 = 0$ and for all $i \geq 1$, $t_i - t_{i-1} = \Delta > 0$.

Property:

The sequence $(S^{i,0})_{i\geq 1}$ is stationary and ergodic.

$\rightarrow$ **Illustration:**



Usefull tools : Cdf Supremum

Let $(X_t)_{t\geq 0}$ be a stationary OU process with parameter $\theta = (a, b, \beta)$, the cdf of the random variable $S^{1,0} = sup_{s\leq\Delta} X_s$ is :

$$\mathbb{P}(S^{1,0} < c) = -\frac{e^{-(c-\frac{a}{b})^2\frac{b}{\beta}}}{\sqrt{2\pi}} \sum_{n\geq 1} e^{-b\mu_{n,c,\theta}\Delta} \frac{D_{\mu_{n,c,\theta}-1}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)}{\mu_{n,c,\theta}\partial_\mu D_{\mu_{n,c,\theta}}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)},$$

with $\mu_{n,c,\theta}$ the positive (ordered) zeros of the function $\mu \mapsto D_\mu\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)$ and $D_\mu(.)$ the Parabolic Cylinder function.

Remark

The cdf of the supremum is closely related with the first hitting time probability over a constant boundary of the OU process.

First insight

For a sample $(S^{1,0}, \cdots, S^{N,0})$, we introduce the likelihood :

$$L_N(\theta) = f(S^{1,0}, \cdots, S^{N,0}, \theta),$$

with f the joint probability density of $(S^{1,0}, \cdots, S^{N,0})$, and the estimator is :

$$\tilde{\theta}_N = (\tilde{a}_N, \tilde{b}_N, \tilde{\beta}_N) = \mathrm{Argmax}_{\theta \in \Theta} L_N(\theta),$$

with $\Theta \subset \mathbb{R} \times \mathbb{R}_+^\star \times \mathbb{R}_+^\star$.

Issue : The law of the variable $(S^{1,0}, \cdots, S^{N,0})$ is unknown and complicated to compute.

Idea :

$$L_N(\theta) \approx \mathcal{L}_N(\theta) = \prod_{i=1}^{N} f_\Delta(S^{i,0}, \theta),$$

with f_Δ the probability density of $S^{1,0}$.

1. A pseudo-likelihood estimator of the Ornstein-Uhlenbeck parameters from suprema observations

ρ-mixing property

For any $n, h \in \mathbb{N}^\star$, and for f and g in an appropriate class of measurable functions:

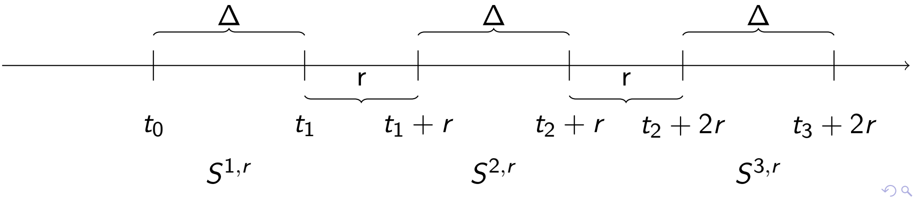
$$\rho(h) = \mathrm{Corr}\left(f\left(S^{1,0}, \cdots, S^{n,0}\right), g\left(S^{n+h,0}, \cdots\right)\right) \xrightarrow[h \to +\infty]{} 0.$$

Mixing consequence

For $r \in \mathbb{R}_+$, we have $S^{i,r} = sup_{s \in [t_{i-1}+(i-1)r, t_i+(i-1)r]} X_s$ and

$$\mathrm{Corr}\left(f\left(S^{n,r}\right), g\left(S^{n+1,r}\right)\right) \xrightarrow[r \to +\infty]{} 0.$$

Furthermore, $(S^{i,r})_{i \geq 1}$ is stationary and ρ-mixing (also ergodic).



1. A pseudo-likelihood estimator of the Ornstein-Uhlenbeck parameters from suprema observations

Estimation procedure

We denote $\theta_0 = (a_0, b_0, \beta_0) \in \Theta$, the parameter to be estimated and $\mathbb{P}_{\theta_0}$ the measure of supremum's observation.

Let $r \in \mathbb{R}_+$, $N \in \mathbb{N}^\star$, for a sample $(S^{1,r}, \cdots, S^{N,r})$, we introduce the following pseudo-likelihood function :

$$\mathcal{L}_N^r(\theta) = \prod_{i=1}^{N} f_\Delta(S^{i,r}, \theta),$$

with f_Δ the probability density of $S^{1,r}$. The estimator is :

$$\hat{\theta}_N = (\hat{a}_N, \hat{b}_N, \hat{\beta}_N) = \underset{\theta \in \Theta}{\mathrm{Argmax}} \mathcal{L}_N^r(\theta).$$

Estimator property

1. Identifiability: For $\theta_1, \theta_2 \in \Theta$, $\mathbb{P}_{\theta_1} = \mathbb{P}_{\theta_2} \implies \theta_1 = \theta_2$.

2. The estimator $\hat{\theta}_N$ is weak consistent *i.e.*:

$$\hat{\theta}_N \xrightarrow[N \to +\infty]{\mathbb{P}_{\theta_0}} \theta_0.$$

3. The following convergence is verified:

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow[N \to +\infty]{law} \mathcal{N}_3\left(0, I_{\theta_0}^{-1}\right),$$

with I_{θ_0} Fisher information matrix associated to the parameter θ_0.

Sketch of proof for consistency (one parameter) Part 1

For a sequence of observations $(S^{1,r}, \cdots, S^{N,r})$ with parameters b_0 :

- $\hat{b}_N = \text{Argmax}_{b\in\Theta} M_N(b)$ with :

$$M_N(b) = \frac{1}{N}\sum_{i=1}^{N} \log\left(\frac{f_\Delta(S^{i,r}, b)}{f_\Delta(S^{i,r}, b_0)}\right).$$

- Ergodic Theorem :

$$M_N(b) \xrightarrow[N\to+\infty]{a.s} M(b) = \mathbb{E}_{b_0} \log\left(\frac{f_\Delta(., b)}{f_\Delta(., b_0)}\right).$$

- The supremum's law is identifiable then $\text{Argmax}_{b\in\Theta} M(b) = b_0$

- Using the compactness of Θ and some domination property, we prove that :

$$\mathbb{P}_{b_0}\left(\sup_{\Theta}|M_N - M| \geq \epsilon\right) \xrightarrow[n\to+\infty]{} 0.$$

Then :

$$\hat{b}_N \xrightarrow[N\to+\infty]{\mathbb{P}_{b_0}} b_0.$$

The domination property is obtain with **asymptotic expansion on the density**, closely related to the μ-zeros of Parabolic Cylinder functions.

Trade off between the number of observations N_{tot} and the gap r

Let $r = (k-1)\Delta$, with $k \in \mathbb{N}^\star$ and N_{tot} the number of observations (for $r = 0$). What is the optimal gap r ?

$\rightarrow$ The coefficient r influences the independency in our observations.

$\rightarrow$ The number of observations N_{tot} influences the efficiency of our estimator.

- The idea is to bound :

$$\text{Mean Squared error of } \mathcal{L}^r_{\lceil N_{tot}/k \rceil} \leq g(\Delta, N_{tot}, k, \theta_0).$$

- The appropriate gap for the estimation is $r^\star = (k^\star - 1)\Delta$ with

$$k^* = \underset{1 \leq k \leq N}{\text{Argmin}} \; g(\Delta, N_{tot}, k, \theta_0).$$

We use the set (Nb of observations, gap) $= (\lceil N_{tot}/k^\star \rceil, r^\star)$.

Simulated Data

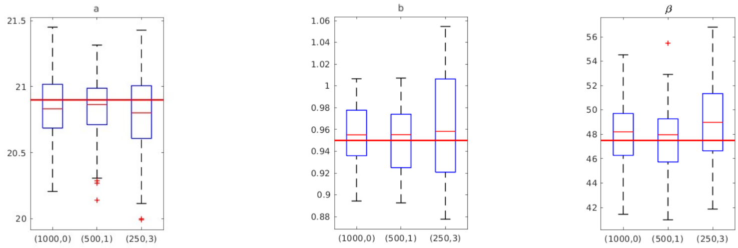
We denote (N, r), the set of numerical parameters, with N the number of suprema observations and r the gap between this observations.

We simulate a stationary OU $(X_t)_{0 \leq t \leq T}$ with parameter $\theta_0 = (a_0, b_0, \beta_0) = (20.9, 0.95, 47.5)$ using an Euler scheme with $T = 10^3$ days and $dt = 10^{-3}$.

Suprema observations are taken over time windows with length $\Delta = 1$ day.

We apply our estimation method on 100 simulated trajectories for three different set of numerical parameters, $(1000, 0)$, $(500, 1)$ and $(250, 3)$.

The optimal set of numerical parameters is $(500, 1)$.



(a) (b) (c)

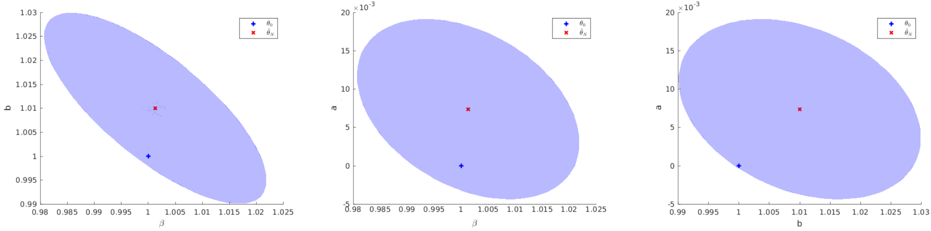
Figure – Estimation of the parameters $(a_0, b_0, \beta_0) = (20.9, 0.95, 47.5)$ of an OU process, the red line corresponds to the theoretical value of the parameter.

Numerical parameters	Relative RMSE	ME
(250, 3)	(0.0154, 0.0615, 0.0701)	(-0.1072, -0.0063, 0.6946)
(500, 1)	(0.0109, 0.0351, 0.0557)	(-0.0538, -0.0096, 1.0821)
(1000, 0)	(0.0113, 0.0348, 0.0578)	(-0.0482, -0.0074, 1.0693)

Table – Table of the relative RMSE and ME for the estimator of $\theta_0 = (a_0, b_0, \beta_0) = (20.9, 0.95, 47.5)$ with different numerical parameters.

95% Confidence Ellipsoïd

Using the Central Limit Theorem, we draw the 95% Confidence Ellispoïd.



(a) βb plane (b) βa plane (c) ba plane

Figure – Cut planes of the 95% confidence ellispoïd associate to the estimator of $\theta_0 = (20.9, 0.95, 47.5)$ and the set of numerical parameters $(N, r) = (500, 1)$.

Estimation Issues

Estimation issues arise :

- Overestimation/bias on the β estimator comes from the decrease of $\beta \mapsto \mathcal{L}^r_N(\mu, \lambda, \beta)$.
- For $\beta >> b$, the β estimator will have a big variation.
- Better results can be obtained by fixing β and performing the estimation method on parameters a and b (2D-estimation).

Weather data

We use the data set of Paris's temperature measurements. It records only maximum, minimum and mean daily temperature from 1900 to nowadays. In our application, we study daily summer temperature.

- From 15th of june to the 14th of august (61 days) each year between 1950 and 1984 included (2135 days).
- We take a gap of $r = 1$ day between each observation.

Comparaison of results

We therefore apply our estimation protocol on these data, we obtain $\hat{\theta} = (\hat{a}, \hat{b}, \hat{\beta}) = (17.076, 0.841, 36.306)$.

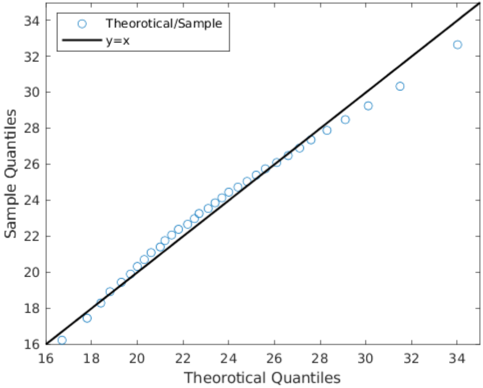


Figure – QQ-plot for the estimation on weather data-set.

Conclusion

Work carried out :

- Optimal gap r between observations.
- Application to a weather data-set.

Opening :

- Estimate the parameters of an OU process using a sample of suprema and infima observations. Law of the supremum and infimum $\rightarrow$ related with the first exit time of an interval of an OU process.

2. Parameters estimation of a Threshold CKLS process from continuous and discrete observations

Introduction on Threshold process :

Definition : A threshold process is a stochastic process where the dynamics change when the process crosses specific threshold levels.

Example : Oscillating Brownian motion

$$dX_t = (\sigma_+ \mathbb{1}_{X_t > 0} + \sigma_- \mathbb{1}_{X_t < 0})\, dB_t, \quad t \geq 0$$

- Closely related to Skew diffusion.
- Threshold autoregressive (TAR) models in discrete time were introduced in the early 1980s.
- Applications in finance, physics and meteorology.

Introduction : CKLS include many famous special cases

Geometric Brownian motion/Black-Scholes model,
OU process/Vasicek model, CIR process... are special cases of

Standard Chan-Karolyi-Longstaff-Sanders (CKLS) process

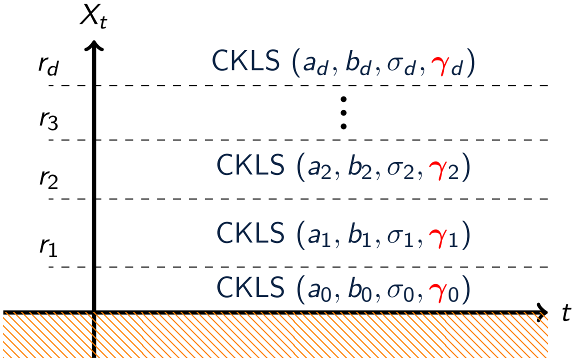
$$X_t = x_0 + \int_0^t (a - bX_s)ds + \int_0^t \sigma X_s^{\gamma} dB_s, \quad t \geq 0,$$

with $x_0, a, \sigma > 0$, $b \in \mathbb{R}$, $\gamma \in \{0\} \cup [1/2, 1]$.

Focus $\gamma = 1/2$: the Cox–Ingersoll–Ross (CIR) process is 1D, with continuous non negative trajectories and exhibits mean-reverting behavior.

- In finance, it may describe the evolution of interest rates.
- In biology, it may appear in population dynamics.
- Linked with Bessel and Square-Bessel processes.

We consider **Threshold CKLS**: different CKLS on different space intervals.



On the SDE: This just translates in piecewise constant coefficients.

T-CKLS as an SDE

Let $(X_t)_{t\geq 0}$, the process solution of

$$X_t = X_0 + \int_0^t a_d(X_s) - b_d(X_s)X_s ds + \int_0^t \sigma_d(X_s)(X_s)^{\gamma_d(X_s)} dB_s, \quad t \geq 0,$$

Example (d thresholds, $d \in \mathbb{N}$):

$$a_d(x) = \sum_{j=0}^{d} a_j \mathbb{1}_{I_j}(x)$$

with $I_j = [r_j, r_{j+1})$ for all $j \in \{1, \cdots, d\}$ and $I_0 = (0, r_1)/[0, r_1)/(-\infty, r_1)$, where $r_0 = 0 < r_1 < \cdots < r_d < r_{d+1} = +\infty$, $d \in \mathbb{N}$.

$\rightarrow$ Threshold-Geometric Brownian motion ($\gamma = 1$), Threshold-OU process ($\gamma = 0$), Threshold-CIR process ($\gamma = 1/2$)...

Content

Context : Observe a unique trajectory of the T-CKLS process.

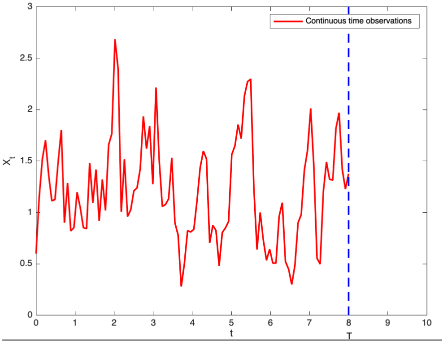
Joint estimation of the parameter $(a_j, b_j, \sigma_j)_{j=0}^{d}$, using two diffrents contrast functions.

- MLE and QMLE for $(a_j, b_j)_{j=0}^{d}$.
- Quadratic estimator for $(\sigma_j)_{j=0}^{d}$.

$\rightarrow$ Study of the asymptotic behavior of the estimators (consistency, asymptotic normality).

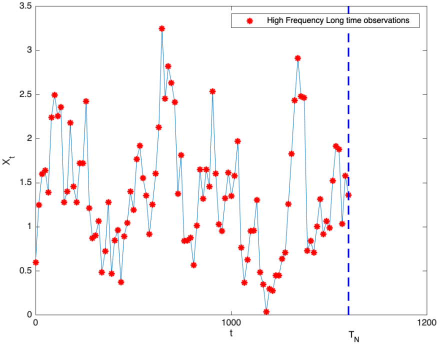
In the case :

- Continuous time observations (ergodic regime).
- High frequency and long time observations (ergodic-stationary regime).



Continuous time observations:

We observe the process on the interval $[0, T]$ with $T \in (0, \infty)$.



High Frequency in Long time observations:

We observe the process on the discrete time grid $0 < t_0 < \cdots < t_N = T_N$, for $N \in \mathbb{N}^{\star}$ and $T_N \in (0, \infty)$.

Estimation method : Continuous time observations

Assume that we observe the entire trajectory $(X_t)_{t\in[0,T]}$.

We can define the Likelihood (Girsanov weights) is :

$$\mathcal{L}_T(a,b) := \exp\left(\int_0^T \frac{a_d(X_s)-b_d(X_s)X_s}{\sigma_d(X_s)^2(X_s)^{2\gamma_d(X_s)}}\,dX_s - \frac{1}{2}\int_0^T \frac{(a_d(X_s)-b_d(X_s)X_s)^2}{\sigma_d(X_s)^2(X_s)^{2\gamma_d(X_s)}}\,ds\right).$$

The MLE is $(\hat{a}_T,\hat{b}_T) = \underset{(a,b)\in\Theta}{\mathrm{Argmax}}\ \mathcal{L}_T(a,b)$ (explicitly computable).

$\rightarrow$ Consistency and asymptotic normality when $T\rightarrow\infty$.

Estimation method : Discrete time observations

We observe the process on a time grid $t_0 < \cdots < t_{N-1} < T_N < \infty$. Where $X_i := X_{t_i}$, with $i = 0, \cdots, N$ and $\Delta_N = \max_{k=0\cdots,N-1} \{t_{k+1} - t_k\}$.

We do not have a true MLE because the law of the T-CKLS is not explicit.

Consider the discretized likelihood

$$\text{disc-}\mathcal{L}_{T_N,N}(a,b) = \exp\left(\sum_{i=0}^{N-1} \frac{a_d(X_i) - b_d(X_i)X_i}{2\sigma_d(X_i)X_i}(X_{i+1} - X_i) - \frac{t_{i+1} - t_i}{4} \frac{(a_d(X_i) - b_d(X_i)X_i)^2}{\sigma_d(X_i)X_i}\right).$$

The estimator is then $(a_{T_N,N}, b_{T_N,N}) = \underset{\theta\in\Theta}{\text{Argmax}}\ \text{disc-}\mathcal{L}_{T_N,N}(\theta)$.

$\rightarrow$ Consistency and asymptotic normality when $T_N \rightarrow \infty$ and $\Delta_N \rightarrow 0$.

2. Parameters estimation of a Threshold CKLS process from continuous and discrete observations

Applications :

- Thresholds parameter $(r_j)_{j=1}^d$ by using the likelihood or the quasi-likelihood.
- The parameter $(\gamma_j)_{j=0}^d$ is difficult to estimate (open problem).
- Testing on the existence of one or more thresholds on the interest rate data [Su & Chan, 2017].

$$(\text{Test}) \begin{cases} H_0 : & \text{Null hypothesis} & (d) \text{ thresholds} \\ H_1 : & \text{Alternative hypothesis} & (d+1) \text{ thresholds} \end{cases}$$

with $d \in \mathbb{N} \rightarrow$ Repeat the test as long as it is significant.

Thank you for your attention.

Bibliography

Blanchet-Scalliet C., Dorobantu D., Nieto B. (2023). Some properties for ν-zeros of Parabolic Cylinder functions. *Le Matematiche.*Volume LXXVIII, pages 277-287.

Blanchet-Scalliet C., Dorobantu D., Nieto B. (2024). A pseudo-likelihood estimator of the Ornstein-Uhlenbeck parameters from suprema observations. *Statistical Inference for Stochastic Processes.* Volume 27, pages 407–425.

Mazzonetto S., Nieto B. (2024). Parameters estimation of a Threshold CKLS process from continuous and discrete observations. *Submitted.*

Blanchet-Scalliet C., Dorobantu D., Nieto B. (2024^{+})**.** On the Threshold killed Ornstein-Uhlenbeck process. *Ungoing work.*

Mazzonetto S., Nieto B. (2024^{+})**.** Strong existence and uniqueness to thresholds SDEs. *Ungoing work.*

Pseudo-Likelihood estimator of the OU parameters :

Alili L., Patie P., Pedersen J.L. (2005). Representations of the First Hitting Time Density of an Ornstein-Uhlenbeck Process *Stochastic Environmental Research and Risk Assessment*

Blanchet-Scalliet C., Dorobantu D., Gay L. (2018). Risk assessment using suprema data *Stochastic Environmental Research and Risk Assessment*

Gobet E., Matulewicz G. (2016). Parameter estimation of Ornstein-Uhlenbeck process generating a stochastic graph *Statistical Inference for Stochastic Processes*

Doukhan P. (1994). Mixing *Springer*

Parameters estimation of a Threshold CKLS :

Ben Alaya M., Kebaier A. (2012). Parameter estimation for the square-root diffusions : ergodic and nonergodic cases *Stochastic Models*

Su, F., Chan, K. S. (2017). Testing for threshold diffusion *Journal of business economic statistics*

Lejay A. and Pigato P. (2018). Statistical estimation of the oscillating Brownian motion *Bernoulli*

Lejay A. and Pigato P. (2020). Maximum likelihood drift estimation for a threshold diffusion *Scandinavian Journal of Statistics*

Mazzonetto S., Pigato P. (2020). Drift estimation of the threshold Ornstein-Uhlenbeck process from continuous and discrete observations *Statistica sinica*

Trade off between r and N

- For N fixed, $r = (k-1)\Delta$ with $k \in 1, N$, $(S^{i,r})_{1 \le i \le \lceil N/k \rceil}$ rather than $(S^{i,0})_{1 \le i \le N}$.
- We have :

$$\mathbb{E}\left[\left(\frac{1}{\lceil N/k \rceil} \log \mathcal{L}^r_{\lceil N/k \rceil} - \frac{1}{\lceil N/k \rceil}\mathbb{E}\left[\log \mathcal{L}^r_{\lceil N/k \rceil}\right]\right)^2\right] \le g(\Delta, N, k, \theta_0).$$

- The appropriate gap for the estimation is $r^\star = (k^\star - 1)\Delta$ with

$$k^* = \underset{1 \le k \le N}{\text{Argmin}}\; g(\Delta, N, k, \theta_0)$$

.

Least square methods in [Blanchet-Scalliet, 2018] (Part 1)

- N observed suprema with Δ the size of the observations
- $N_q \in \mathbb{N}^\star$, $\forall j = 1, ..., N_q, s^{j,0} \in \mathbb{R}$.
- Least square method by minimizing the function :

$$Q_N(\theta) := \sum_{j=1}^{N_q} \left[F\left(s^{j,0}, \theta, \Delta\right) - F_N^*(s^{j,0}) \right]^2 .$$

where F is the cdf of the supremum, F_N^* is the empirical cdf on the

- Parameters estimated by the least square estimator :

$$\tilde{\theta}_N = (\tilde{a}_N, \tilde{b}_N, \tilde{\beta}_N) = \underset{\theta \in \Theta}{\text{Argmin}} \; Q_N(\theta) ,$$

with Θ a compact subset of $\mathbb{R} \times \mathbb{R}_+^\star \times \mathbb{R}_+^\star$.

Least square methods in [Blanchet-Scalliet, 2018] (Part 2)

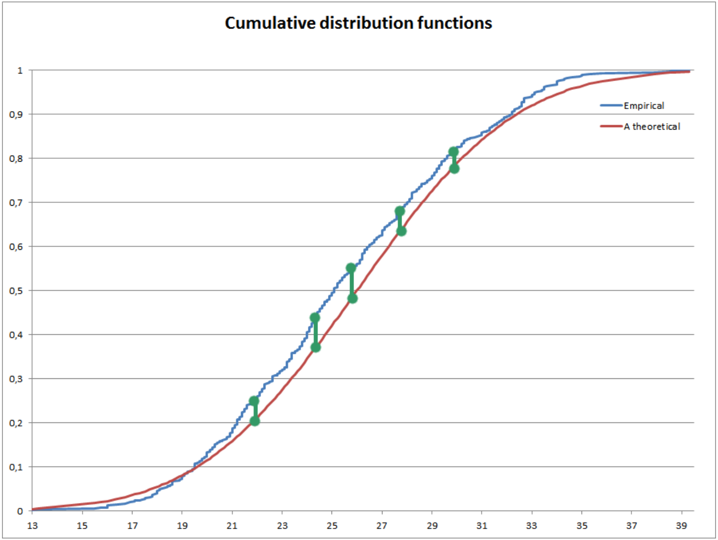


Figure – Least Square Methods

Pseudo likelihood vs Least square (Part 1)

For $c, x_0 \in \mathbb{R}$ such that $x_0 \leq c$, let us note:

$$\tau_c = \inf\{t \geq 0, X_t = c\} \quad \textit{and} \quad \mathbb{P}_{x_0} = \mathbb{P}(.|X_0 = x_0)$$

Link between Supremum and Hitting time

For all $c \in \mathbb{R}$ and $x_0 \in \mathbb{R}$, we have :

$$\mathbb{P}_{x_0}(\tau_c > t) = \mathbb{P}_{x_0}\left(\sup_{s<t} X_s \leq c\right)$$

In [Alili, 2005], they get the hitting time law of the OU.

- Bessel bridge (using Girsanov) $\Longrightarrow$ Integral representation (Least square method).
- Inverting the Laplace transform of τ_c $\Longrightarrow$ Series representation (Pseudo likelihood method).

Pseudo likelihood vs Least square (Part 2)

- Lower numerical cost using the series representation of the supremum Law.
- Better results for the pseudo likelihood method (series form) against the Least square methods (integral form/series form) in terms of RMSE and ME.

Remark 1

Consistency for the least squares method can be obtained by using the same technical assumptions used for the likelihood method.

Remark 2

We can use the series representation in the least squares method and increase the number of quantiles used. However, the pseudo-likelihood method remains slightly better.

Pseudo likelihood vs Least square (part 3)

Least square method (Series representation + 10 quantiles) :

Numerical parameters	Relative RMSE	ME
(250,3)	(0.0241, 0.0863, 0.0722)	(-0.0957, -0.0072, -0.8463)
(500,1)	(0.0126, 0.0359, 0.0572)	(0.0593, 0.0101, -1.0012)
(1000,0)	(0.0135, 0.0375, 0.0594)	(-0.0488, -0.0091, -1.0750)

Pseudo likelihood method :

Numerical parameters	Relative RMSE	ME
(250,3)	(0.0154, 0.0615, 0.0701)	(-0.1072, -0.0063, -0.6946)
(500,1)	(0.0109, 0.0351, 0.0557)	(-0.0538, -0.0096, -1.0821)
(1000,0)	(0.0113, 0.0348, 0.0578)	(-0.0482, -0.0074, -1.0693)

Table – Table of the relative RMSE and ME for the estimator of $\theta_0 = (a_0, b_0, \beta_0) = (20.9, 0.95, 47.5)$ with different numerical parameters.

Joint law between supremum and infimum (Part 1)

We consider an OU process X solution of

$$\begin{cases} dX_t = (a - bX_t)dt + \sqrt{\beta}dB_t, \\ X_0 = x_0, \end{cases}$$

Let $\tau_{[c_1,c_2]}$, the first exit time of an OU process from the interval (c_1, c_2) :

$$\tau_{[c_1,c_2]} = \inf\{t \geq 0, X_t \notin (c_1, c_2)\}.$$

Link between the joint law of the supremum and the infimum with the law of the first exit time $\tau_{[c_1,c_2]}$:

$$\mathbb{P}_{x_0}\left(t \leq \tau_{[c_1,c_2]}\right) = \mathbb{P}_{x_0}\left(c_1 < \inf_{s\leq t} X_s,\ \sup_{s\leq t} X_s < c_2\right).$$

Joint law between supremum and infimum (Part 2)

We denote $\mathbb{P}_{x_0}\left(t \le \tau_{[c_1,c_2]}\right) = p(x_0,t)$ the function p is solution of

$$\begin{cases} \partial_t p(x_0,t) = \frac{\beta}{2}\partial^2_{{x_0}^2} p(x_0,t) + (a - bx_0)\partial_{x_0} p(x_0,t) & (x_0,t) \in (c_1,c_2)\times\mathbb{R}^\star_+, \\ p(c_1,t) = p(c_2,t) = 0 & t \in \mathbb{R}^\star_+, \\ p(x_0,0) = 1 & x_0 \in [c_1,c_2]. \end{cases}$$

- Spectral decomposition method $\implies$ explicit solution for p (cf Chapter 4):

$$p(x_0,t) = \sum_{n=1}^{\infty} c_n(t) e_n(x_0).$$

where $(e_n)_{n\ge 0}$ are the normalized solutions of:

$$\begin{cases} \mathcal{L}f(x_0) = -\mu f(x_0), & \forall x_0 \in [c_1,c_2]. \\ f \in \mathrm{Dom}^W(\mathcal{L}). \end{cases}$$

with $\mathcal{L}f = \frac{\beta}{2}f'' + (a - bx)f'$

- The solution is numerically unstable.